

A METHOD OF PRE-ANALYSIS OF A MACHINE-READABLE FORM IMAGE.

Inventors:

ZUEV, KONSTANTIN, RU;
FILIMONOVA, IRINA, RU;
ZLOBIN, SERGEY, RU.

Current U.S. Class:

Intern'l Class:

G04K009/62

Field of Search:

References Cited U.S. Patent Documents

6169822	January 2, 2001	Jung	382/186
6148119	November 14, 2000	Takaoka	382/203
6137905	October 24, 2000	Takaoka	382/186
5592572	January 7, 1997	Le	382/190
5471549	November 28, 1995	Kurosu et al.	382/195
5235651	August 10, 1993	Nafarieh	382/202
5031225	July 09, 1991	Tochikawa et al.	382/203

References Cited Non-patent documents

1. "Proceedings of the 13-th International Conference on Pattern Recognition, August 25-29, 1996, Vienna, Austria". Vol.III, Track C, IEEE Computer Society Press, Los Alamitos, California, p.681-685.
2. 1. "Proceedings of the 13-th International Conference on Pattern Recognition, August 25-29, 1996, Vienna, Austria". Vol.III, Track C, IEEE Computer Society Press, Los Alamitos, California, p.691-695.
3. "Proceedings of the 13-th International Conference on Pattern Recognition, August 25-29, 1996, Vienna, Austria". Vol.III, Track C, IEEE Computer Society Press, Los Alamitos, California, p.696-700.
4. "Proceedings of the 13-th International Conference on Pattern Recognition, August 25-29, 1996, Vienna, Austria". Vol.III, Track C, IEEE Computer Society Press, Los Alamitos, California, p.701-705.
5. "Proceedings of the 13-th International Conference on Pattern Recognition, August 25-29, 1996, Vienna, Austria". Vol.III, Track C, IEEE Computer Society Press, Los Alamitos, California, p.768-772.
6. "Proceedings of the 13-th International Conference on Pattern Recognition, August 25-29, 1996, Vienna, Austria". Vol.III, Track C, IEEE Computer Society Press, Los Alamitos, California, p.793-797.

A METHOD OF PRE-ANALYSIS OF A MACHINE-READABLE FORM IMAGE.

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates generally to an optical character pre-recognition of machine-readable forms, and in particular to bit-mapped image and one or more model matching methods, and also image spatial direction identification.

Prior Art.

According to widely known methods of text pre-recognition a bit-mapped image is parsed into regions, containing text and/or non-text regions, with the further dividing said text regions into objects, containing strings, words, character groups, characters etc.

Some known methods uses preliminarily document type identification for narrowing a list of possible documents types, examined in an analysis of the document logical structure.

According to this group of methods the document type identification is an independent step of document analysis, forestalling logical structure identification. At that the document type and its properties list become defined up to the moment of defining the logical structure thereof. Or wise versa, a document structure identification may be an integral part of logical structure identification process. In this case the document type that fits closer the analyzed image is selected.

A spatial orientation direction verification is present in a number of documents.

In the US patent #5031225 (July 09, 1991, Tochikawa et al.) is disclosed a method of document image spatial orientation verification, using a preliminarily assigned character, to be found

in the document. The found character is recognized to fit one of the 4 models thereof, corresponding with four possible directions.

The most reliably matching model indicates the orientation direction of the image.

The method causes a mistake in the case of possible different directions of text orientation to be present in the document. It also may cause mistake if the character is not reliably recognized after converting into image state.

In the US patent #5235651 (August 10, 1993, Nafarieh) the orientation direction of the image is estimated via setting up and accepting a hypothesis on the level of initial image units by analyzing the transition from dark points (pixels) and regions to light ones and vice versa. If the examined hypothesis is not accepted, the new one is set up, considering the image to be turned at 90° angle.

The method can't work if various orientation directions of text can be present on the form.

In the US patent #5471549 (November 28, 1995, Kurosu et al.) to define the image orientation direction the text characters are selected from the text one after another and are tried to recognize, supposing orientation direction to be 0°, 90°, 180°, 270°. The direction of the best matching is assumed as the right document image orientation.

The method can't work if various orientation directions of text can be present on the form as in the previous example.

In the US patent #5592572 (January 7, 1997, Le) the problem is solved by dividing the image into a large amount of objects, either of text or non-text types. Then the orientation of all initial objects is estimated via recognition of characters, with the further joining them into large ones and estimating the orientation thereof. Finally there is the only text object, covering the whole text field with the corresponding orientation estimation.

The main shortcoming of the method lies in that the orientation estimation is performed along with recognition of text portions, thus reducing the method output.

In the US patents #6137905 (October 24, 2000, Takaoka) and #6148119 (November 14, 2000, Takaoka) the orientation direction is estimated by dividing the image into a plurality of regions, possessing various estimation weight coefficient. Then the orientation direction is estimated via the text recognition in the said regions. The total direction is estimated as a sum of particular ones together with their weight coefficients.

The shortcoming of the method is the low method output, depending greatly upon the recognition results.

In the US patent #6169822 (January 2, 2001, Jung) the predetermined portion of the text is parsed from the image and is performed (processed) recognition. In the case of recognition failure, the inference is made about the other orientation direction of the image.

To achieve the reliable result via the said method the large number of text portions are to be recognized. That surely reduces the method output.

SUMMARY OF THE INVENTION

One or more objects of the form are assigned thereon, composing graphic image, unambiguously defining its direction of spatial orientation. The said graphic image properties comprise a description of a special model for defining the direction of spatial orientation. Identification of the image with the said model the right direction of image spatial orientation is defined. The said model properties are stored in a special data storage means, one of the embodiment of which is the form image model description.

In the similar way one or more form objects are assigned thereon, composing graphic image, unambiguously defining its type.

Additionally one or more form objects may be assigned, for the case of profound form type analysis, if two or more forms are close in appearance or in properties list. The graphic image properties comprise description of a special model for form type definition. The said model properties are stored in a special data storage means, one of the embodiment of which is a form model description.

After converting the form image is parsed into regions containing text images, data input fields, special reference points, lines and other objects.

The possible distortion, caused by the document conversion to electronic state, is eliminated from the image.

Objects, comprising the graphic image for spatial orientation verification, are identified on the form image. The orientation direction accuracy is verified and corrected if necessary.

The objects, comprising the graphic image for form type definition, are identified on the form image. The proper model is selected via identification of the said graphic image. In the case of multiple identification result, the profound analysis of the form type is performed. The profound analysis is performed in the similar way adding the supplementary objects to the graphic image and performing new identification.

The profound analysis is performed automatically or fully or partly manually.

BRIEF DESCRIPTION OF THE DRAWING

Fig. 1 shows a document with three assigned elements comprising graphic image.

Fig. 2 shows the graphic image, formed by assigned elements.

Fig. 3 shows some examples of graphic objects used as assigned elements comprising graphic image.

DETAILED DESCRIPTION OF THE INVENTION

The document logical structure examination requires dividing the document image into elements of different types. The single element of the document can contain its title, authors name, date of the document or the main text etc.

The composition of the document elements depends upon its type.

The document logical structure is performed by the following ways:

- on the base of fixed elements location,
- using a table or multi-column structure [1], [5], [6],
- on the base of structural images identification [4],
- via specialized methods for special documents types [3].

Methods of the first group requires fixed structural elements location and are used for fields mark out, i.e. image regions, containing elements for documents of standard form [2]. The exact elements location on the form may be distorted by scanning. The distortion may be of various kinds: shift, a small turn angle, a compression and stretching, a large turn angle.

All kinds of distortion are eliminated on the first stage of document image processing.

The coordinates of regions may be founded relatively to:

- image edges,
- special reference points,
- remarkable form elements,
- a correlation function, taking into account of all or a part of the listed above.

Sometimes, the distortion may be ignored due to its negligibility. Then the image coordinates are computed relatively to the image edges.

The most of the methods for form type identification uses special graphic objects that are reliably identified reference points, as black squares or rectangles, a short dividing lines composing cross or corner (Fig.3) etc. Searching the reference

points location combination as an image using the special models, the type of the analyzed form can be correctly defined.

The main technical result of the invention consists in gaining

- universality of the pre-recognition analysis of machine-readable forms,
- ability to process documents' images of more than one form type in one session,
- ability to process documents images of different directions of spatial orientation,
- ability to perform the pre-recognition process with high output.

The said technical result is achieved in the following way.

One or more objects (1) are assigned on the form, composing graphic image (2), unambiguously defining its direction of spatial orientation. The said graphic image properties are described in a special model used for defining the direction of spatial orientation. Identification of the said image via the said model the right direction of image spatial orientation is defined. The said special model properties are stored in a special data storage means, one of the embodiment of which is the form image model description.

In the similar way one or more form objects (1) are assigned, composing graphic image (2) on the form, unambiguously defining its type. Additionally one or more supplementary form objects may be assigned for profound form type analysis, if two or more forms are close in appearance or in properties list. The graphic image properties is described of an another special model used for form type definition. The said another special model properties are stored in a special data storage means, one of the embodiment of which is a form model description.

After converting to electronic state the form image is parsed into regions containing text objects images, data input fields, special reference points, lines and other objects.

The possible distortion, caused by the document conversion to electronic state, is eliminated from the image.

The objects, comprising the graphic image for spatial orientation verification, are identified on the form image. The orientation direction accuracy is verified and corrected if necessary.

The objects, comprising the graphic image for form type definition, are identified on the form image. The matching model is selected via identification of the said graphic image. In the case of multiple identification result, the profound analysis of the form type is performed. The profound analysis comprises creation of a new special model for form type identification containing primary special model plus one or more supplementary form objects. The image is performed a supplementary identification using new special model.

The profound analysis may be performed fully or partly automatically.

One or more form object, comprising the graphic image is described in a form of alternative.

References Cited Non-patent documents

1. "Proceedings of the 13-th International Conference on Pattern Recognition, August 25-29, 1996, Vienna, Austria". Vol.III, Track C, IEEE Computer Society Press, Los Alamitos, California, p.681-685.
2. 1."Proceedings of the 13-th International Conference on Pattern Recognition, August 25-29, 1996, Vienna, Austria". Vol.III, Track C, IEEE Computer Society Press, Los Alamitos, California, p.691-695.
3. "Proceedings of the 13-th International Conference on Pattern Recognition, August 25-29, 1996, Vienna, Austria". Vol.III, Track C, IEEE Computer Society Press, Los Alamitos, California, p.696-700.
4. "Proceedings of the 13-th International Conference on Pattern Recognition, August 25-29, 1996, Vienna, Austria". Vol.III, Track C, IEEE Computer Society Press, Los Alamitos, California, p.701-705.
- 5."Proceedings of the 13-th International Conference on Pattern Recognition, August 25-29, 1996, Vienna, Austria". Vol.III, Track C, IEEE Computer Society Press, Los Alamitos, California, p.768-772.
- 6."Proceedings of the 13-th International Conference on Pattern Recognition, August 25-29, 1996, Vienna, Austria". Vol.III, Track C, IEEE Computer Society Press, Los Alamitos, California, p.793-797.